

Using GRACE to estimate snowfall accumulation and assess gauge undercatch corrections in high latitudes

**Ali Behrangi^{1,2}, Alex Gardner², John T. Reager², Joshua B. Fisher², Daqing Yang³,
George J. Huffman⁴, Robert F. Adler⁵**

¹University of Arizona, Department of hydrology and atmospheric sciences, Tucson, AZ,
USA.

²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.

³National Hydrology Research Center, Environment Canada, Saskatoon, Saskatchewan,
Canada

⁴NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

⁵Earth System Science Interdisciplinary Center, University of Maryland, College Park,
Maryland, USA

Accepted in: *Journal of Climate*, 13 August 2018

Corresponding author:

Ali Behrangi,
1133 E. James E Rogers Way, Harshbarger Rm. 226B
Tucson, AZ 85721-0011
E-mail: Behrangi@email.arizona.edu

Abstract

Ten years of terrestrial water storage anomalies from the Gravity Recovery and Climate Experiment (GRACE) were used to estimate high latitude snowfall accumulation using a mass balance approach. The estimates were used to assess two common gauge-undercatch correction factors (CFs): Legates climatology (CF-L) utilized in the Global Precipitation Climatology Project (GPCP), and Fuchs dynamic correction model (CF-F) used in the Global Precipitation Climatology Centre (GPCC) Monitoring product. The two CFs can be different by more than 50%. CF-L tended to exceed CF-F over northern Asia and Eurasia, while the opposite was observed over North America. Estimates of snowfall from GPCP, GPCC-L (GPCC corrected by CF-L), and GPCC-F (GPCC corrected by CF-F) were 62%, 64%, and 46% more than GPCC over northern Asia and Eurasia. GRACE-based estimate (49% more than GPCC) was the closest to GPCC-F. We found that as near surface air temperature decreases, the products increasingly underestimated the GRACE-based snowfall accumulation. Overall, GRACE showed that CFs are effective in improving GPCC estimates. Furthermore, our case studies and overall statistics suggest that CF-F is likely more effective than CF-L in most of the high latitude regions studied here. GPCP showed generally better skill than GPCC-L, which might be related to the use of satellite data or additional quality controls on gauge inputs to GPCP. This study suggests that GPCP can be improved if it employs CF-L instead of CF-F to correct for gauge undercatch. However, this implementation requires further studies, region-specific analysis, and operational considerations.

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

1. Introduction

Multiple studies have shown significant changes associated with a warming climate in high latitude environments (Moritz et al. 2002; Smith et al. 2005; Alley et al. 2007; Ye et al. 2014, 2016) and evidence for future amplifications of these changes (Solomon et al. 2007; Lau et al. 2013; Serreze and Francis 2006). Increasing temperatures can enhance moisture transport from low to high latitudes (Stephens et al. 2016) that could significantly influence Arctic precipitation amount, type and timing. Changes in high-latitude water and energy budgets can create feedbacks to weather and climate systems and influence regional and global carbon cycles. Reported changes in the high latitude hydrologic cycle are numerous, including changing precipitation and water vapor (Ye et al. 2016), snow cover, mountain glaciers and ice sheets, lake size (Smith et al. 2005; Shepard et al. 2018; Gardner et al. 2013), groundwater storage (Zhang et al. 2017), and freshwater discharge (e.g., Yang et al. 2002; McClelland et al. 2006), all impacted by precipitation amount and its distribution. Critical to understanding these changes and the interactions between system components is a high-quality data record. Among various datasets, precipitation is fundamental, as it often directly affects other variables, such as snow cover and river flow. However, accurate quantification of precipitation over cold regions has been a longstanding challenge (Serreze and Barry 2005; Behrangi et al. 2016; Wen et al. 2017). The work in this paper directly addresses the accuracy of the observational precipitation datasets on which climate modelers, hydrologists, glaciologists, ecologists, biologists, engineers and others depend to advance understanding of the hydrological cycle at high latitudes.

A major source of global precipitation data are ground stations that often provide the longest records, sometimes spanning several decades. However, precipitation measurements from gauges can contain large errors in high latitudes for two primary reasons: (1) the precipitation gauge network is sparse and discontinuous in most regions—a substantial decrease in the number of high latitude precipitation stations since 1990 has exacerbated this limitation; and, (2) precipitation measurement often must be bias corrected to account for wetting loss and gauge undercatch (Goodison et al. 1998; Yang et al. 2001). The bias correction is largest for solid precipitation, can be as large as 300% (Fuchs et al. 2001), and depends on the choice of correction method. Furthermore, different countries use different gauge and windshield combinations that introduce additional complexity to the bias correction process (Scaff et al. 2015). There are valuable efforts to study the consistency of precipitation observations between various manual and automatic gauges, among these are efforts under the World Meteorological Organization (WMO) Solid Precipitation Intercomparison Experiment (SPICE) project (Rasmussen et al., 2012) that have investigated many automated instruments for snowfall observations in various climate conditions. The data collection from SPICE, and the subsequent analyses, can enhance our ability to measure snowfall and improve bias correction methods. However, most of these efforts have not yet been fully implemented by the most widely-used precipitation products (e.g., GPCC and GPCP).

Station measurements are often used to construct gridded products at different spatial and temporal resolutions. An example is the gridded daily and monthly precipitation products constructed by the Global Precipitation Climatology Centre (GPCC). GPCC has been used in numerous studies and products such as the Global Precipitation Climatology Project (GPCP).

Such gridded datasets facilitate evaluation of model and remote sensing products. However, in high latitudes such gridded estimates are greatly limited by distant interpolation between sparse observations.

Over the last few decades space-borne sensors have provided another major source for estimating precipitation with global or near-global coverage (over both land and ocean). Continuous efforts have been dedicated to improve the quality of precipitation measuring sensors and retrieval techniques. However, high latitude precipitation estimation has remained a major area of uncertainty (Adler et al. 2012; Behrangi et al. 2014a,b 2016) mainly due to insufficient sensitivity of sensors to light rain and snowfall, poor understanding of polar precipitation microphysics, difficulties in separating light rain from clouds (Berg et al. 2006; Lebsock and L'Ecuyer 2011), and sensitivity of retrievals to prior knowledge about precipitation phase (Liu, 2008). These uncertainties may be disproportionately larger over land due to orographic precipitation enhancement (Shige et al. 2013) and unknown surface emissivity (Ferraro et al. 2013), especially over frozen land. These challenges have largely limited measurements of precipitation in high latitudes from space, especially during cold months when snowfall is typical.

With the launch of the CloudSat Cloud Profiling Radar (CPR) in 2006 (Stephens et al. 2008), major advancements in the quantification of high latitude precipitation have been achieved. These advancements mainly come from the high sensitivity of CloudSat to drizzle, light rain, and snowfall that are the dominant types of precipitation in high latitudes. However, CloudSat has poor temporal coverage (i.e., 16-day repeat orbit and less temporal coverage after 2011 due to

battery issues). This means that CloudSat often misses precipitation events between two satellite overpasses. While CloudSat has been effective in estimating total precipitation over high latitude oceans (Behrangi et al. 2014), two major issues have limited its application over land: (1) precipitation is estimated at 1.00 - 1.50 km above the surface to avoid contamination of the reflectivity profile by surface returns (Tanelli et al. 2008) that can result in missing shallow precipitation; and, (2) over land, only the snowfall product is officially operational, while the rainfall is produced in an experimental mode that has not been evaluated.

The operation of the Global Precipitation Measurement (GPM) mission (Hou et al. 2014; Skofronick-Jackson et al. 2017) starting in 2014 has also advanced retrieval of light rain and snowfall relative to its predecessor, the Tropical Rainfall Measurement Mission (TRMM). Measurement advancements come from the higher sensitivity of the GPM core instruments to light rain, snowfall, and improvements in retrieval methods. Unfortunately, the GPM Core Observatory instruments do not cover regions poleward of $\sim 65^{\circ}\text{S/N}$, where retrieval of precipitation has to rely on other sensors and remains highly uncertain.

In this study, we show that NASA's Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage anomaly data provide information that can be used to advance the development of high latitude precipitation datasets. Precipitation accumulation is calculated from GRACE Terrestrial Water Storage Anomalies (TWSA) using the mass conservation principle (e.g. Dingman, 2008) discussed in Section 2. This concept has been used to study various components of the water cycle including the near surface terrestrial water storage (e.g., snow, soil moisture, and glaciers) and groundwater (e.g., Niu et al. 2007; Gardner et al. 2013). GRACE

has also been used for precipitation analyses. Swenson (2010) used GRACE to assess cold-season precipitation from the GPCP and Climate Prediction Center Merged Analysis of Precipitation (CMAP) and found that while spatial patterns and inter-annual variability are highly correlated between the datasets, differences typically increase at higher latitudes. Focusing on the East Antarctic Ice Sheet, Boening et al. (2012) investigated the balance between mass loss primarily due to glacial discharge and mass gain through precipitation derived from reanalysis, CloudSat, and GRACE. Recently, Behrangi et al. (2017) used GRACE to estimate monthly precipitation over high mountain Asia, where other precipitation datasets face major shortcomings.

This study is different from previous investigations mainly because: (1) for the first time, GRACE is used to assess the gauge undercatch corrections that otherwise may not be possible through observations; (2) we use recent advancements in GRACE data processing (Watkins et al. 2015) that result in increased accuracy and reduced bias due to leakage effects; and, (3) we minimize the contribution of uncertainties from model-derived evapotranspiration and runoff estimates (that are typically needed for component-wise mass balance) by focusing on frozen surfaces. As gauge precipitation measurement is a key component of gauge-only (e.g., GPCC) or satellite-gauge products (e.g., GPCP and similar products), this analysis is important to produce and assess high latitude precipitation products.

The present study benefits from the unique capability of GRACE that enables an independent estimate of precipitation accumulation (i.e., gravimetry versus radiometry with no need for empirical parameterizations and ground-based calibration). Furthermore, GRACE measures total

accumulation so does not miss precipitation events in terms of total distributed water between two satellite overpasses. As such, the cumulative nature of GRACE estimates also mitigates errors that can be caused due to missing light rain, snow and mixed-phase precipitation by other space-borne sensors (Behrangi et al. 2012; 2014c).

2. Datasets

2.1. GPCC products

GPCC was established in 1989 under WMO, in part to provide the in situ component of GPCP, a precipitation product of the Global Energy and Water Exchanges (GEWEX) project. By integrating station data from several networks, GPCC produces station-based gridded precipitation data that covers global land excluding the Antarctic (Schneider et al. 2014). To produce a complete map, GPCC has to rely on interpolation of sparse in situ data that can lead to large and poorly characterized errors. This is especially the case in high latitudes. For example, only a handful of gauges operate over the entire Antarctic, so few that GPCC chooses not to provide estimates for this continent. GPCC offers various products including the two widely-used products that are assessed in this study: (1) the Full data monthly (V7), currently available from 1901 through 2013, which incorporates all stations and goes through the highest quality-control process, but is not available in real-time and does not correct for gauge under-catch; and, (2) the low latency monthly Monitoring product (V5) that uses records from fewer gauging stations and is available from 1982 through to present. In this paper GPCC refers to GPCC Full data monthly (V7), unless we explicitly mention GPCC Monitoring. The V5 product also provides additional information including the fraction of liquid and solid precipitation and an

empirical gauge correction factor (CF). The CF is based on a method described in Fuchs et al. (2001). In brief, the method uses a dynamic correction model (Forland et al. 1996; Rubel and Hantel 1999) and employs various synoptic observations such as air temperature, relative humidity, precipitation intensity and wind speed. Daily bias-correction of synoptic observations are performed after determining precipitation phase using air and dew point temperatures, originally collected from more than 6×10^5 global synoptic data from 16 winter months. Occasionally, missing information limits the application of this method. In such cases the Legates climatology (Legates and Willmott 1990) can be used (see Fuchs et al. 2001 for further details). The Legates correction method is a monthly spatially varying climatology of undercatch estimate using the best guess of gauge types around the world as of the early 1980's. Legates and Willmott (1990) used this method to generate global climatology of monthly precipitation by removing systematic errors caused by wind, wetting on the interior walls of the gauge, and evaporation from the gauge observations.

2.2. GPCP

GPCP is a widely-used community-based analysis of global precipitation under the auspices of the World Climate Research Program (WCRP) and GEWEX. GPCP uses a combination of spaceborne sensors over land and ocean including the Special Sensor Microwave Imager (SSM/I), Special Sensor Microwave Imager/Sounder (SSMIS), geostationary imagers and polar orbiting infrared sounders such as the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) and Atmospheric Infrared Sounder (AIRS). Gauge-based estimates are adjusted for gauge undercatch using the Legates CFs described earlier. Gauge measurements are averaged with satellite estimates using inverse estimated error variance

weighting (Huffman et al. 1997). Several upgrades have been made to the GPCP since its initial release (Huffman et al. 1997; Adler et al. 2003; Huffman et al. 2009). The latest version of GPCP (V2.3) (Adler et al. 2016, 2017) was released in late 2016 at monthly $2.5^{\circ} \times 2.5^{\circ}$ resolution, covering the period 1979 to present and is available from http://eagle1.umd.edu/GPCP_ICDR/. As described in Adler et al. (2016), the changes in GPCP V2.3 are generally small compared to the previous version (GPCP V 2.2), and are mainly related to the updates in satellite cross-calibration procedures and new sets of gauge analyses used in GPCC V7 Full analysis. Monthly GPCP V2.3 was used in our study here.

2.3. GRACE

GRACE is a twin satellite mission that utilizes radar interferometry to measure range-rate variations between the two satellites, and has been operational since 2002. The GRACE orbit is perturbed by changes in gravity due to mass redistribution, such as water, within the Earth System. Orbital perturbations are recorded as changes in range-rates that can then be used to invert for the surface mass variations with high accuracy (Tapley et al. 2004). While several factors can result in mass variations, at monthly scale mass variations are largely caused by redistribution of water at or near the Earth's surface. We use the latest 1° scaled version of the JPL GRACE TWSA mascon data product (RL05M version 2). This version includes improvements in both background geophysical models and orbital parameterization, and is based on the recent 'mascon' solution' of surface mass change (Watkins et al. 2015), reducing solutions uncertainty and bias due to leakage effects. By using land surface model a priori information on spatial variability, a 0.5° signal is recovered as described in Landerer and

Swenson (2012). This dataset is available since 2003 and can be obtained from GRACE Tellus website (available at <https://grace.jpl.nasa.gov/>).

2.4. CloudSat

CloudSat 94-GHz nadir-looking CPR, with a spatial resolution of $\sim 1.4 \times 1.7$ km, flies in NASA's A-train constellation and crosses the equator at approximately 01:31/13:31 local times (Stephens et al. 2008). CloudSat has separate products that provide estimates of rain and snowfall rates. The snowfall product, the 2C-SNOW-PROFILE (Wood et al., 2014), is available over both land and ocean and has been evaluated by various groups (e.g., Smalley et al. 2014; Chen et al. 2016). However, CloudSat rain estimates (Haynes et al. 2009; Lebsock and L'Ecuyer 2011) are currently limited to ocean. Over land, the estimates are provisional with limited evaluation. As this study focuses on high latitude land, only 2C-SNOW-PROFILE product is used. The product is available via the CloudSat data processing center: <http://www.cloudsat.cira.colostate.edu/data-products>.

2.5. Other datasets

A few other variables are also used in this work, including two-meter near surface air temperature (T2m), evapotranspiration (ET), runoff, and sublimation. T2m was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA-Interim) (Dee et al. 2011), ET and runoff were obtained from the Global Land Data Assimilation System (GLDAS; (Rodell et al. 2004) using Noah and VIC Land Surface Models. The products provide monthly data at $0.25^\circ \times 0.25^\circ$ resolution (GLDAS_NOAH025_M, GLDAS_VIC025_M - V2.0) and are available from NASA Goddard Earth Sciences (GES) Data

and Information Services Center (DISC). Sublimation was obtained from the latest Modern-Era Retrospective Analysis (MERRA2-land) (Rienecker et al. 2011; Bosilovich et al. 2015) sublimation product ('EVPSBLN'), available at monthly $1/2^\circ \times 2/3^\circ$ resolution from DISC.

The products listed above were mapped onto common monthly $2.5^\circ \times 2.5^\circ$ grids prior to analysis. This is mainly because CF-L and the latest version of GPCP (V2.3) were only available at monthly $2.5^\circ \times 2.5^\circ$ resolution. This resolution allows for the collection of large number of CloudSat samples within individual $2.5^\circ \times 2.5^\circ$ degree cells, which is important for stability of the derived statistics (Behrangi et al. 2012, 2016). The period of study is from 2003 to 2013, the former set by GRACE and the latter by GPCC monitoring product data availability.

3. Method and results

3.1. Precipitation accumulation from GRACE

GRACE TWSA is used to calculate precipitation accumulation based on the mass conservation principle (e.g., Dingman 2008). If t_1 represents accumulation start-time (e.g., first day of a month) and t_2 represents accumulation end-time (e.g., last day of a month) for domain D, the water storage change between time t_1 and t_2 can be calculated as:

$$\Delta S = \int_{t_1}^{t_2} P(t)dt - \int_{t_1}^{t_2} ET(t)dt - \int_{t_1}^{t_2} Sub(t)dt - \int_{t_1}^{t_2} Q_{net}(t)dt \quad (\text{Eq. 1})$$

Where ΔS is the change in storage (i.e., from GRACE) between time t_1 and t_2 and $P(t)$, $ET(t)$, $Sub(t)$, $Q_{net}(t)$ are the precipitation rate, evapotranspiration rate, sublimation rate, and net lateral flux rate (e.g., runoff) for domain D at time t , respectively.

252

253 Because the objective of this study is to investigate the use of GRACE as an alternate tool for
254 estimating high latitude precipitation, it is important to reduce uncertainties in various
255 components of Eq. 1 (i.e., runoff, sublimation, and evapotranspiration). Therefore, in the present
256 study, Eq. 1 was applied only to frozen land, defined by grids with individual 6-hour T2m less
257 than a Temperature Threshold (TT) of 273K for the entire time during boreal winter (defined
258 here as December, January, and February; hereafter DJF). Mean DJF temperature and areas
259 corresponding to five selected TTs are shown in Figure 1. While the TT constraint limits the
260 study area, it can largely reduce estimation uncertainties. By implementing this constraint: (1)
261 the role and uncertainties associated with Q_{net} and ET are minimized; and, (2) the focus will be
262 on snowfall rate estimate, which continues to be uncertain from both satellite (Behrangi et al.
263 2014b) and gauge (Strangeways 2006) measurements. Figure 2 shows that CFs are often larger
264 for snowfall than rainfall measurements, making the final snowfall rate estimate highly
265 dependent on the accuracy of the CFs. While the temperature constraints allow analysis of
266 Greenland and Antarctica, they were not studied here mainly because: (1) the use of GRACE to
267 estimate the accumulated precipitation in Greenland and Antarctic ice sheets is complicated by
268 ice divergence (continual export of mass from the interior of the ice sheets to the oceans via ice
269 flow; Gardner et al. 2018); and, (2) GPCC products have no coverage over Antarctica. To
270 minimize assumptions related to Q_{net} , ET, and sublimation amount, these components were
271 included in Eq. 1 even though they are generally small. Most cold-season water budget studies
272 do not consider sublimation, but here it is considered because we found that sublimation can
273 account for as much as 7% of the total water budget over some regions and can be larger than the

total contribution of ET and runoff that, on average, account for less than 3% of total precipitation during winter.

3.2. Analysis of gauge correction factor (CF)

CFs are generally much larger for snowfall than rainfall (Figure 2). This makes it critical to identify a proper CF for snowfall analysis. Figure 3 shows seasonal maps of two CFs: Legates climatology (hereafter CF-L) and Fuchs dynamic correction model (hereafter CF-F). CF-L consists of fixed monthly climatology factors based on Legates and Willmott (1990) and was obtained from GPCP through communications between GPCP and GPCC. As described in Section 2.1, CF-F is dynamic in time and obtained from the GPCC monthly Monitoring product. In Figure 3, maps of CF-F are calculated by averaging a decade (2000-2009) of monthly CFs for each season. CF-L covers global land, but CF-F does not cover Antarctica. The maps in Figure 3 show that CFs are close to (but above) one near equator, and are larger in higher latitudes and over mountainous areas. CFs are much larger during winter than summer in high latitudes, reflecting the dominance of snowfall there in the winter (Fig. 2). Both CF-L and CF-F agree on the zonal and seasonal dependence of the CFs, but they show major regional differences (e.g., compare the CFs over Eurasia in winter) as also noted by Schneider et al. (2017).

A comparison of CFs is provided in Figure 4, which shows seasonal maps of percent relative difference between CF-F and CF-L. The relative difference is calculated by subtracting CF-F from CF-L and dividing the result by the mean of the two using a decade (years 2000-2009) of monthly data. The comparison shows that the relative differences are maximum during winter (e.g., can exceed 50%) and minimum during summer, confirming that CFs are more uncertain for

snowfall. CF-L values mostly exceed CF-F over Europe and northwest and northeast Asia, but are smaller over North America and part of East Asia (Figures 4a, 4b, and 4d). Observed differences can be related to the methodology used in determining CFs. Different gauges have different specifications such as different outsplash, evaporative and aerodynamic characteristics, and errors. It is often not an easy task to correct for these, and the detailed metadata required are not usually available (Strangeways 2006). Adjustment for errors for snowfall can be even more difficult due to the aerodynamic effects. As discussed by Strangeways (2006), such corrections are not straightforward and further efforts and considerations are needed to improve them. The present work can be a valuable approach to assess CFs by investigating which CF shows a better match to the total accumulation of snowfall captured by GRACE.

3.3. Comparison of precipitation products

Given that CF-F and CF-L can have significant regional and seasonal differences, it is insightful to apply the CFs to GPCC and inter-compare the outcomes with other precipitation estimates. Figure 5 shows winter and summer mean precipitation from GPCP (5a and 5b) and different GPCC products (5c-5h) calculated from a decade of data (2003-2012). By multiplying CF-L and CF-F by GPCC Full data (Figs 5g and h), GPCC-L (Figs 5c and 5d) and GPCC-F (Figs 5e and 5f) are determined and used for cross-comparison with GPCP and GPCC. GPCP covers both ocean and land, and thus provides a complete view of precipitation pattern. Over land, GPCP uses GPCC Full data and CF-L to correct gauge measurements before combining them with satellite estimate. Therefore, it can be seen that GPCP and GPCC-L are similar over land. However, GPCP has its own quality control method and does not use all GPCC information (Huffman et al. 1997). By comparing the precipitation products in Figure 5, it can be concluded

that the largest discrepancy among the products occur during winter, when the choice of CF can significantly influence the estimated precipitation patterns. This is especially noticeable over Eurasia and North Asia, where GPCC originally shows much less precipitation than what estimated from GPCC-L and GPCC-F. The fact that the GPCC-L and GPCC-F can be very different from GPCC is highly important, as in many studies, gauge measurement (e.g., from GPCC) is considered as an absolute truth, yet with no undercatch correction.

We compare total accumulation of snowfall from GPCC, GPCC-L, GPCC-F, GPCP, and CloudSat with the GRACE estimate (using Eq. 1). We focus on northeastern and northcentral Asia, representing a large area of frozen land (e.g., see Fig. 1b) to reduce GRACE-based precipitation estimation uncertainties as discussed in Section 3.1. Furthermore, this region has a wide range of elevations and temperatures, allowing detailed comparison of different products as a function of elevation and temperature conditions. The inclusion of CloudSat in our analysis is especially useful, as CloudSat is the most sensitive sensor in space to estimate snowfall rate since late 2006. However, CloudSat observations are only available during daylight hours after 2011 (due to battery problem) and more than a month worth of data was lost in winter of 2010. Therefore, in the present work our CloudSat analysis is limited to 2007-2009.

Mean winter snowfall rates from various products are shown in Figures 6a-f. The maps are constructed at $2.5^{\circ} \times 2.5^{\circ}$ resolution using two years of data (2007-2008) to reduce potential noise. This is especially important to improve CloudSat sampling. Corresponding maps of surface elevation above sea level (hereafter elevation) and station counts are shown in Figures 6g and 6h, respectively. The high elevation regions are mainly located in the southern part of the study area

with extension to the northeast (Fig. 6g). It can also be seen that the stations are sparse, especially in the north, where most of the grids contain either one or no station (Fig. 6h). Figures 6a-f show that the products generally agree on the spatial distribution of winter snowfall, but they also display large regional differences in snowfall rates. GPCP (Fig. 6b) and GPCC-L (Fig. 6e) agree well on both rate and spatial distribution of snowfall. They also show the largest snowfall rates over the northwest part of the study area compared to the other products. The observed overestimation by GPCP and GPCC-L over this region is also in agreement with Behrangi et al. (2014c), in which precipitation was estimated from atmospheric water vapor sinks calculated from various observations and reanalyses. On the other hand, GPCC (Fig. 6c) and CloudSat (Fig 6f) show good agreement, but both display the lowest snowfall rates among all other products. As GPCC does not incorporate CFs, it can be concluded that CloudSat likely underestimate the true seasonal snowfall amount. Compared to GPCC-L, and consistent with Figures 3 and 4, GPCC-F (Fig. 6d) has lower snowfall rates over the northwest and slightly higher rates over the central and eastern part of the study area. This flatter distribution in GPCC-F agrees more with the GRACE estimate (Fig 6a) than other products.

Snowfall rates are shown in Figure 7 and are constructed from the maps displayed in Figure 6. The GRACE estimate is shown on the abscissa and estimates from other products are shown on the ordinate. Summary statistics are also provided in each panel. GPCP (Fig. 7a), GPCC-F (Fig. 7c), and GPCC-L (Fig. 7d) show the highest Correlation Coefficient (CC) with GRACE. GPCC-F shows the lowest Root Mean Square Error (RMSE) and closest relative BIAS compared to GRACE, suggesting that GPCC-F is likely the best match to GRACE. Relative BIAS is calculated by dividing the products snowfall rate by the corresponding rate from GRACE.

Therefore, a product with a relative bias of one has the same average as GRACE. Among the products studied, GPCC (Fig. 7b) and CloudSat (Fig 7e) show lower agreement with GRACE. While it is not surprising to see GPCC, which does not include CFs, shows lower skill than GPCC-L and GPCC-F, it remains to be investigated why CloudSat demonstrates low skill scores. However, as discussed earlier, CloudSat precipitation is estimated in the fifth range bin (1.0 - 1.5 km) above the surface to avoid contamination of the reflectivity profile by surface returns (Tanelli et al. 2008; Smalley 2014). Our analyses using three years (2007-2009) of CloudSat Lidar cloud classification product (2B-CLDCLASS-LIDAR-R04; available from the CloudSat data processing center) over the study region (displayed in Fig 6) shows that on average about 26% of precipitable clouds reside below 1.5 km from the surface. Here precipitable clouds are defined as cloud types that are expected to precipitate (Sassen et al. 2008; Wang 2011): Stratocumulus (Sc), Cumulus (Cu), Nimbostratus (Ns), and Deep convective (DC). This observation is close to the ~30% underestimation of CloudSat seasonal snowfall relative to GRACE, GPCP, and corrected GPCC products.

While Figure 6 provides valuable insights on regional differences among the studied products, it is also important to relate the differences to other factors that might impact snowfall rate and characteristics. Such analyses can help diagnose reasons behind the observed discrepancies among products and provide more useful feedback to algorithm developers. Figure 8 compares mean snowfall rates as a function of TT (Fig. 8a) and elevation (Fig. 8b). Figure 8 is constructed from the maps displayed in Figure 6. For $TT \geq 270$ there is good agreement among GRACE, GPCP, GPCC-F, and GPCC-L, suggesting that the CFs are effective in improving agreement between GPCC and GRACE estimates. As TT decreases, the products show smaller mean

389 snowfall rates and tend to have larger underestimation compared to GRACE. For TT = 255K, the
 390 underestimation is about 16%, 20%, and 27% for GPCC-F, GPCP, and GPCC-L, respectively.
 391 GPCC and CloudSat produce comparable mean snowfall rates across the entire range of TT (Fig
 392 8a), but both products show relatively large underestimations compared to other products. For
 393 example, GPCC underestimates GPCC-L by about 27%, GPCC-F by about 36%, and GRACE by
 394 about ~46% for TT = 255K. Figure 8b shows a similar analysis by replacing TT with surface
 395 elevation. For elevations below 200m, where seasonal mean snowfall is often largest (see Fig. 6),
 396 GRACE and GPCC-F show almost identical mean snowfall rates. Their estimates are about 20%
 397 less than that estimated by GPCC-L and GPCP, but about 25% and 50% more than CloudSat and
 398 GPCC, respectively. As elevation increases, all products show lower snowfall rates. This is
 399 expected as elevation and T2m are correlated. CloudSat consistently shows higher estimates than
 400 GPCC, but lower estimates than other products across the entire elevation ranges. At high
 401 elevation ranges (i.e., elevation > 1000m), CloudSat shows a snowfall rate similar to GPCC-F,
 402 GPCC-L, and GPCP, but about 50% underestimation compared to GRACE. Analysis of
 403 CloudSat-Lidar cloud classes (2B-CLDCLASS-lidar ; Sassen et al. 2008) for 2007-2009 shows
 404 that about 40% of precipitable cloud classes can be missed at elevations greater than 1000 m.
 405 This is obtained by calculating the fraction of precipitating clouds (out of total precipitating
 406 clouds) that fall in the bottom part of the profile, defined by surface elevation plus 1.5km surface
 407 contamination height. The missed 40% ratio supports the likelihood that CloudSat is
 408 underestimating about 50% at high elevations. Note that 2B-CLDCLASS-lidar combines
 409 CloudSat CPR and CALIPSO lidar measurements to classify clouds into eight groups: stratus
 410 (St), stratus (Sc), cumulus (Cu, including cumulus congestus), nimbostratus (Ns), altocumulus
 411 (Ac), altostratus (As), deep convective (cumulonimbus), or high (cirrus and cirrostratus) clouds.

Among these classes, Sc, Cu, Nb, and deep convective are considered to have precipitation rate greater than zero in their definition (Wang et al. 2013).

Time series of winter mean snowfall using a decade of the studied products over frozen regions in north and northeast Asia are compared in Figure 9,. The mean snowfall rates are based on a two-year running average to reduce potential uncertainties. The GRACE estimate is not shown for winter of 2011, when more than a month of GRACE observations were unavailable due to battery management. All products display a generally similar pattern to GPCC, but with a considerably larger snowfall rate (Fig. 9). On average GPCP, GPCC-L, and GPCC-F show about 62%, 64%, and 46% higher snowfall rates than GPCC. GRACE estimate shows about 49% higher rate than GPCC and falls between GPCC-L and GPCC-F estimates, but closer to GPCC-F.

Skill scores for quantitative comparison between GRACE and other products in estimating winter cumulative snowfall for 2003-2013 (except 2011) are provided in Table 1. GPCP and GPCC-F are found to have the higher and GPCC has the lower Pearson Correlation Coefficient (CC) with GRACE. Furthermore, GPCC-F shows the lowest RMSE and the closest relative bias compared to GRACE, suggesting that GPCC-F outperforms other products according to GRACE. It can also be seen that GPCP has slightly better skill than GPCC-L, which might be related to the use of remotely sensed information or further quality controls on gauge datasets by GPCP.

434 Figure 10 compares quantitative statistics (CC, relative RMSE, and relative bias) calculated
 435 between GRACE winter-snowfall rates and other products as a function of TT (top row) and
 436 elevation (bottom row) using a decade of corresponding datasets (2004-2013). Because snowfall
 437 rate tends to be small at lower temperature (or higher elevations), in Figure 10 we show relative
 438 RMSE, calculated by dividing RMSE values by GRACE snowfall rate. Figure 10a shows that as
 439 TT decreases, the correlation between products and GRACE estimate decreases substantially.
 440 GPCP and GPCC-F show the largest and GPCC have the smallest correlation with GRACE
 441 across most temperature zones. Similar results can be seen by replacing TT with elevation (Fig
 442 10d), although in this case GPCC-F outperforms GPCC-L (in matching GRACE) only at
 443 elevations below 350m and above that GPCP and GPCC-L outperform GPCC-F. With respect to
 444 relative RMSE versus TT (Fig 10b), no clear conclusion can be made, except a monotonic
 445 increase in relative RMSE of GPCC as TT decreases. However, Figure 10e shows that the
 446 products generally show higher relative RMSE at higher elevations. Figures 10b and 10e show
 447 that GPCC-F has almost equal or lower relative RMSE than other products, across most TT or
 448 elevation ranges. Figure 10c shows that as temperature threshold decreases, the products display
 449 a larger decrease in mean snowfall rate than GRACE, resulting in a monotonic decrease in
 450 relative BIAS. This is also the case as elevation increases (Fig 10f). GPCC-F has the closest
 451 mean snowfall rate to GRACE for $TT > 270K$ and $TT < 260$, but for TT between 260K and
 452 270K, GPCC-L and GPCP show higher agreement with GRACE (Fig 10c). Similarly, for
 453 elevations below 350m, GPCC-F shows the closet mean snowfall rate to GRACE. At higher
 454 elevations, all products (except GPCC) converge, but underestimate relative to GRACE (Fig.
 455 10f). The estimates from GPCP, GPCC-L, and GPCC-F are as large as GRACE for elevations

between 350m and 750m (Fig. 10f). GPCC consistently underestimates GRACE, as large as ~35% to 45% for different TTs and 23% to 63% for different elevations (Figs 10c and 10f).

Thus far, the analyses were mainly performed over the north and northeastern Asia. However, as shown in Figure 1b, similar analysis can also be done over northwestern part of North America continent. On average, this region contains only about 16% of total grids in the northern Asia region and presents a narrow range of TT (see Fig. 1b). Therefore, satisfactory analysis using TT or elevation was not found useful, mainly due to insufficient sample counts. Figure 11 is similar to Figure 9, but is constructed over northwestern part of North America continent using TT = 272. Figure 11 shows that GPCP and GPCC-F have closer mean snowfall rates to GRACE than other products. Despite the results in northern Asia, Figure 11 shows that GPCC-L estimate is much less than GPCP. This observation, together with further analysis using spatial maps, suggest that combination of satellite and gauge analysis used in GPCP has likely increased GPCP estimate, thus it is closer to GPCC-F than GPCC-L. CloudSat estimates for 2008 and 2009 are also closer to GPCC-F than GPCC-L. This suggests that from GRACE and CloudSat perspective, GPCC-F might produce more reliable winter snowfall estimate than GPCC-L over northwestern region of North American continent.

Summary and Concluding Remarks

Accurate quantification of precipitation in high latitudes, and especially during the cold season, has been a longstanding challenge. Precipitation errors can affect a wide range of applications, such as those describing environmental changes associated with climate warming, intensification of hydrologic cycle, changes in regional hydrology and ecosystem, and water and energy budget

calculations (Tian et al., 2007; Ye et al., 2012). In this study, we show that GRACE can add valuable insights on quantification and assessment of high latitude snowfall rates.

Precipitation measurements by gauges are often considered as truth in many studies and provide a key component of gauge-only (e.g., GPCC) or merged satellite-gauge products (e.g., GPCP and similar products). However, precipitation estimates from gauges can be especially erroneous in high latitudes, where the precipitation gauge network is often sparse and gauge undercatch is significant (Yang et al., 2005; Pan et al., 2016). In addition, automatic precipitation gauges have increasingly been used in the northern regions, so the transformation from manual to automated observation can add uncertainties on climate monitoring. Therefore, further efforts to relate and compare the manual and automated precipitation observations (e.g., Groisman et al., 1999; Rasmussen et al., 2012) are needed.

Here, we utilized GRACE to estimate winter snowfall and assess gauge-undercatch correction skill based on two popular correction factors (CFs): Legates climatology (CF-L) used in GPCP, and Fuchs dynamic correction model (CF-F) computed in the latest GPCC Monitoring products. We found that the two correction factors can be significantly different (by more than 50%), especially for snowfall in high latitudes. CF-L tends to show larger values than CF-F over northern Asia and Eurasia, while the opposite is observed over northwestern North America. We performed our analysis using 11 years (2003-2013) of corresponding data sets for grids with surface temperatures continuously below 273K during winter season (DJF). While this constraint limited our precipitation analyses to snowfall and over certain regions, it helped reduce

uncertainties in the ET and runoff components required for closure of the hydrological mass budget.

Our analysis concentrated on the northcentral and northeastern parts of Asia (including most of the Siberia) and compared GRACE estimates of winter snowfall with GPCC, GPCC-F, GPCP-L, GPCP, and CloudSat. “GPCC-F” and “GPCC-L” represent undercatch-corrected GPCC using CF-F and CF-L, respectively. The analysis was also performed at a range of surface elevations and near-surface air temperatures (T2m). Using a decade of data (2003-2013, except 2011), it was found that estimates of winter snowfall from GPCP, GPCC-L, and GPCC-F are about 62%, 64%, and 46% higher than GPCC. The corresponding estimate from GRACE is also about 49% higher than GPCC and falls between GPCC-L and GPCC-F estimates, but closer to GPCC-F. GPCC-F also showed the smallest RMSE and the second highest correlation coefficient (after GPCP) relative to the GRACE estimate. As the temperature threshold decreases (or elevation increases), the products tend to show increasing underestimates compared to the GRACE estimates. Among the products, GPCC-F showed the lowest RMSE compared to GRACE for most of the temperature and elevation ranges studied. We also performed analysis over grids in the northwestern part of North America that showed that GPCP and GPCC-F have mean snowfall rates closer to GRACE than the other products. If one trusts the GRACE approach, GRACE suggests that both CF-L and CF-F significantly improve GPCC estimates, confirming that analyses of original gauge values (including GPCC products) have to be corrected for gauge undercatch prior to use. Additionally, based on case studies and overall statistics, it can be inferred that CF-F is more effective than CF-L in most of the high latitude regions studied here, but regional exceptions also exist.

524

525 The GRACE-based assessment enables an independent estimate of precipitation accumulation
526 (i.e., gravimetry versus radiometry, with no need for empirical parameterizations and ground-
527 based calibration and bias correction). This feature of GRACE makes it especially useful over
528 land in high latitudes, where unknown surface emissivity (particularly over frozen land),
529 sparseness of ground measurement, and orographic precipitation enhancement have long been
530 major challenges for the remote sensing of precipitation. Furthermore, the cumulative nature of
531 GRACE estimates mitigates errors caused by missing precipitation type (e.g., light rain, snow,
532 and mixed-phase) or precipitation events between two satellite overpasses. Another trait that
533 makes GRACE-based estimates of cumulative precipitation, and especially snowfall, more
534 valuable in high latitudes is a reduced dependence on other variables (i.e., ET and runoff) in the
535 mass balance calculations.

536

537 The present approach suggests an alternative means for the estimation of precipitation amount in
538 high latitudes and can be valuable for constraining cumulative precipitation estimates from other
539 methods or for assessing gauge correction coefficients, both of which have been difficult tasks in
540 cold regions. The launch of the GRACE Follow-on mission (gracefo.jpl.nasa.gov) in May of
541 2018 ensures continuation of GRACE-type measurements for several more years, making
542 GRACE-based methods suitable and more useful to assess the long-term precipitation data
543 record.

544

545 **Acknowledgments.** Datasets were collected from various sources: The latest monthly GPCP SG
546 (V2.3) from http://eagle1.umd.edu/GPCP_ICDR/, GPCC Full Data Reanalysis Version 7.0 at

547 1°×1° resolution from ftp://ftp.dwd.de/pub/data/gpcc/html/download_gate.html, GRACE total
548 water storage from <http://grace.jpl.nasa.gov>, and CloudSat products from CloudSat data
549 processing center: <http://www.cloudsat.cira.colostate.edu/data-products>. The research described
550 in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology,
551 under a contract with the National Aeronautics and Space Administration. Financial support was
552 also made available from NASA GRACE and GRACE-FO (NNH15ZDA001N –GRACE) and
553 NASA Energy and Water Cycle Study (NNH13ZDA001N-NEWS), and NASA MEaSUREs
554 (NNH17ZDA001N-MEASURES) awards. Government sponsorship is acknowledged.

Reference:

- Adler, R. F., and Coauthors, 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *Journal of Hydrometeorology*, **4**, 1147-1167.
- Adler, R. F., G. Gu, and G. J. Huffman, 2012: Estimating Climatological Bias Errors for the Global Precipitation Climatology Project (GPCP). *Journal of Applied Meteorology and Climatology*, **51**, 84-99.
- Adler, R. F., M. Sapiiano, G. J. Huffman, D. Bolvin, et al., 2016: New Global Precipitation Climatology Project Monthly Analysis Product Corrects Satellite Data Shifts. *GEWEX News*, **26(4)**, 7-9.
- Adler, R. F., G. Gu, M. Sapiiano, J.-J. Wang, and G. J. Huffman, 2017: Global Precipitation: Means, Variations and Trends During the Satellite Era (1979–2014). *Surveys in Geophysics*, **38**, 679-699.
- Alley, R., T. Berntsen, N. L. Bindoff, Z. Chen, A. Chidthaisong, et al. 2007, Summary for policymakers, in *Climate Change 2007: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the IPCC, edited by S. Solomon et al., pp. 1– 18, Cambridge Univ. Press, Cambridge, U. K.
- Behrangi, A., M. Lebsock, S. Wong, and B. Lambriksen, 2012: On the quantification of oceanic rainfall using spaceborne sensors. *J. Geophys. Res.*, **117**, D20105.
- Behrangi, A., G. Stephens, R. F. Adler, G. J. Huffman, B. Lambriksen, and M. Lebsock, 2014a: An Update on the Oceanic Precipitation Rate and Its Zonal Distribution in Light of Advanced Observations from Space. *Journal of Climate*, **27**, 3957-3965.

- Behrangi, A., Y. Tian, B. H. Lambrigtsen, and G. L. Stephens, 2014b: What does CloudSat reveal about global land precipitation detection by other spaceborne sensors? *Water Resources Research*, **50**, 4893-4905.
- Behrangi, A., S. Wong, K. Mallick, and J. B. Fisher, 2014c: On the net surface water exchange rate estimated from remote-sensing observation and reanalysis. *International Journal of Remote Sensing*, 2170-2185.
- Behrangi, A., and Coauthors, 2016: Status of high-latitude precipitation estimates from observations and reanalyses. *Journal of Geophysical Research: Atmospheres*, **121**, 4468-4486.
- Behrangi, A., A. S. Gardner, J. T. Reager, and J. B. C. G. L. Fisher, 2017: Using GRACE to constrain precipitation amount over cold mountainous basins. *Geophys. Res. Lett.*, doi:10.1002/2016GL071832.
- Berg, W., T. L'Ecuyer, and C. Kummerow, 2006: Rainfall Climate Regimes: The Relationship of Regional TRMM Rainfall Biases to the Environment. *Journal of Applied Meteorology and Climatology*, **45**, 434-454.
- Boening, C., M. Lebsock, F. Landerer, and G. Stephens, 2012: Snowfall-driven mass change on the East Antarctic ice sheet. *Geophys. Res. Lett.*, **39**, L21501.
- Bosilovich, M. G., S. Akella, L. Coy, R. Cullather, and e. al., 2015: MERRA-2: Initial Evaluation of the Climate, NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2015-104606. *National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA*, **43**, 139pp.

- Chen, S., and Coauthors, 2016: Comparison of Snowfall estimates from the NASA CloudSat Cloud Profiling Radar and NOAA/NSSL Multi-Radar Multi-Sensor System. *Journal of Hydrology*, **541**, 862-872.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553-597.
- Dingman, S. L., 2008: *Physical Hydrology*, 2nd ed., Waveland Press, Long Grove, IL, 646pp.
- Ferraro, R. R., and Coauthors, 2013: An Evaluation of Microwave Land Surface Emissivities Over the Continental United States to Benefit GPM-Era Precipitation Algorithms. *Ieee Transactions on Geoscience and Remote Sensing*, **51**, 378-398.
- Forland, E. J., N. m. institutt, and N. W. G. o. P. 1996: *Manual for Operational Correction of Nordic Precipitation Data*. Norske meteorologiske institutt.
- Fuchs, T., J. Rapp, F. Rubel, and B. Rudolf, 2001: Correction of synoptic precipitation observations due to systematic measuring errors with special regard to precipitation phases. *Phys. Chem. Earth*, B29, 689–693.
- Gardner, A. S., and Coauthors, 2013: A Reconciled Estimate of Glacier Contributions to Sea Level Rise: 2003 to 2009. *Science*, **340**, 852-857.
- Gardner, A. S., G. Moholdt, T. Scambos, M. Fahnestock, S. Ligtenberg, M. van den Broeke, and J. Nilsson, 2018: Increased West Antarctic and unchanged East Antarctic ice discharge over the last 7 years. *The Cryosphere*, **12**, 521-547.
- Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison, Rept. 67, 212 pp., World Meteorological Organization, Geneva.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, **95**, 701-722.

- Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bulletin of the American Meteorological Society*, **78**, 5-20.
- Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global precipitation record: GPCP Version 2.1. *Geophys. Res. Lett.*, **36**, L17808.
- Landerer, F. W., and S. C. Swenson, 2012: Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, **48**, W04531.
- Lau, W. K. M., H. T. Wu, and K. M. Kim, 2013: A canonical response of precipitation characteristics to global warming from CMIP5 models. *Geophys. Res. Lett.*, **40**, 3163-3169.
- Lebsock, M. D., and T. S. L'Ecuyer, 2011: The retrieval of warm rain from CloudSat. *J. Geophys. Res.*, **116**, D20209.
- Legates, D. R., and C. J. Willmott, 1990: Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation. *Internat. J. Climatol.*, **10**, 111-127.
- Liu, G., 2008: Deriving snow cloud characteristics from CloudSat observations. *J. Geophys. Res.*, **113**, D00A09.
- McClelland, J.W., S. J. De'ry, B. J. Peterson, R. M. Holmes, and E. F. Wood, 2006: A Pan-Arctic evaluation of changes in river discharge during the latter half of the 20th century, *Geophys. Res. Lett.*, **33**, L06715, doi:10.1029/2006GL025753.
- Moritz, R. E., C. M. Bitz, and E. J. Steig, 2002: Dynamics of Recent Climate Change in the Arctic. *Science*, **297**, 1497-1502.
- Niu, G.-Y., K.-W. Seo, Z.-L. Yang, C. Wilson, H. Su, J. Chen, and M. Rodell, 2007: Retrieving snow mass from GRACE terrestrial water storage change with a land surface model. *Geophys. Res. Lett.*, **34**, L15704.

- Pan, X., and Coauthors, 2016: Bias corrections of precipitation measurements across experimental sites in different ecoclimatic regions of western Canada. *The Cryosphere*, **10**, 2347-2360.
- Rasmussen, R., and Coauthors, 2012: How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed. *Bulletin of the American Meteorological Society*, **93**, 811-829.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, **24**, 3624-3648.
- Rignot, E., J. Mouginot, and B. Scheuchl, 2011: Ice Flow of the Antarctic Ice Sheet. *Science*, **333**, 1427-1430.
- Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, **85**, 381-394.
- Rubel, F., and M. Hantel, 1999: Correction of Daily Rain Gauge Measurements in the Baltic Sea Drainage Basin. *Hydrology Research*, **30**, 191-208.
- Sassen, K., Z. Wang, and D. Liu, 2008: Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *J. Geophys. Res.*, **113**, D00A12.
- Scaff, L., D. Yang, Y. Li, and E. Mekis, 2015: Inconsistency in precipitation measurements across the Alaska–Yukon border. *The Cryosphere*, **9**, 2417-2428.
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, **115**, 15-40.

- Schneider, U., P. Finger, A. Meyer-Christoffer, E. Rustemeier, M. Ziese, and A. Becker, 2017: Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, **8**.
- Serreze, M. C., and R. G. Barry, 2005: The Arctic Climate System, 1st ed., 385 pp., Cambridge University Press, Cambridge
- Serreze, M., and J. Francis, 2006: The Arctic Amplification Debate. *Climatic Change*, **76**, 241-264.
- Shepherd, A., and Coauthors, 2018: Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature*, **558**, 219-222.
- Shige, S., S. Kida, H. Ashiwake, T. Kubota, and K. Aonashi, 2013: Improvement of TMI Rain Retrievals in Mountainous Areas. *Journal of Applied Meteorology and Climatology*, **52**, 242-254.
- Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) Mission for Science and Society. *Bulletin of the American Meteorological Society*, **98**, 1679-1695.
- Smith, L. C., Y. Sheng, G. M. MacDonald, and L. D. Hinzman, 2005: Disappearing Arctic lakes, *Science*, 308(5727), 1429, doi:10.1126/ science.1108142.
- Smalley, M., T. L'Ecuyer, M. Lebsock, and J. Haynes, 2014: A Comparison of Precipitation Occurrence from the NCEP Stage IV QPE Product and the CloudSat Cloud Profiling Radar. *Journal of Hydrometeorology*, **15**, 444-458.
- Solomon, S., et al. 2007, Technical summary, in Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the

- Intergovernmental Panel on Climate Change, edited by S. Solomon et al., pp. 20–91, Cambridge Univ. Press, Cambridge, U. K.
- Stephens, G. L., and Coauthors, 2008: CloudSat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research: Atmospheres*, **113**, D00A18.
- Stephens, G. L., M. Z. Hakuba, M. Hawcroft, J. M. Haywood, A. Behrangi, J. E. Kay, and P. J. Webster, 2016: The Curious Nature of the Hemispheric Symmetry of the Earth's Water and Energy Balances. *Current Climate Change Reports*, 1-13.
- Strangeways, I., 2006: *Precipitation: Theory, Measurement and Distribution*. Cambridge University Press, 302 pp.
- Swenson, S., 2010: Assessing High-Latitude Winter Precipitation from Global Precipitation Analyses Using GRACE. *Journal of Hydrometeorology*, **11**, 405-420.
- Tanelli, S., and Coauthors, 2008: CloudSat's Cloud Profiling Radar After Two Years in Orbit: Performance, Calibration, and Processing. *Geoscience and Remote Sensing, IEEE Transactions on*, **46**, 3560-3573.
- Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins, 2004: GRACE Measurements of Mass Variability in the Earth System. *Science*, **305**, 503-505.
- Tian, X., A. Dai, D. Yang, and Z. Xie, 2007: Effects of precipitation-bias corrections on surface hydrology over northern latitudes. *Journal of Geophysical Research: Atmospheres*, **112**, pp 10.
- Wang, Z., 2011, Process description and interface control document for Level 2 Combined Radar and Lidar Cloud Scenario Classification, available at:
http://www.cloudsat.cira.colostate.edu/sites/default/files/products/files/2B-CLDCLASS-LIDAR_PDICD.P_R04.20120522.pdf.

- Watkins, M. M., D. N. Wiese, D.-N. Yuan, C. Boening, and F. W. C. J. B. Landerer, 2015: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, **120**, 2648-2671.
- Wood, N. B., T. S. L'Ecuyer, A. J. Heymsfield, G. L. Stephens, D. R. Hudak, and P. Rodriguez, 2014: Estimating snow microphysical properties using collocated multisensor observations. *Journal of Geophysical Research: Atmospheres*, **119**, 8941-8961.
- Wang, Z. and coauthors, 2013: Process Description and Interface Control Document 2B-CLDCLASS-LIDAR P_R04 PDICD, available at:
http://www.cloudsat.cira.colostate.edu/sites/default/files/products/files/2B-CLDCLASS-LIDAR_PDICD.P_R04.20120522.pdf
- Wen, Y., P. Kirstetter, J. J. Gourley, Y. Hong, et al. , 2017: Evaluation of MRMS snowfall products over the western United States. *Journal of Hydrometeorology*, **0**, doi: 10.1175/JHM-D-1116-0266.1171
- Yang, D., B. Goodison, J. Metcalfe, P. Louie, E. Elomaa, C. Hanson, V. Golubev, T. Gunther, J. Milkovic, and M. Lapin, 2001: Compatibility evaluation of national precipitation gage measurements, *J. Geophys. Res.*, 106(D2), 1481–1492.
- Yang, D., D. Kane, K. D. Hinzman, X. Zhang, T. Zhang, and H. Ye, 2002: Siberian Lena river hydrological regime and recent change, *J. Geophys. Res.*, 107(D23), 4694, doi:10.1029/2002JD002542.
- Yang, D., D. Kane, Z. Zhang, D. Legates, and B. Goodison, 2005: Bias corrections of long-term (1973–2004) daily precipitation data over the northern regions. *Geophys. Res. Lett.*, **32**, L19501.

- Ye, B., D. Yang, and L. Ma, 2012: Effect of precipitation bias correction on water budget calculation in Upper Yellow River, China. *Environ. Res. Lett.*, **7**, 025201.
- Ye, H., E. J. Fetzer, S. Wong, A. Behrangi, E. T. Olsen, J. Cohen, B. H. Lambriksen, and L. Chen, 2014: Impact of increased water vapor on precipitation efficiency over northern Eurasia. *Geophys. Res. Lett.*, **41**, 2014GL059830.
- Ye, H., E. J. Fetzer, A. Behrangi, S. Wong, B. H. Lambriksen, C. Y. Wang, J. Cohen, and B. L. Gamelin, 2016: Increasing Daily Precipitation Intensity Associated with Warmer Air Temperatures over Northern Eurasia. *Journal of Climate*, **29**, 623-636.
- Zhang, G., and Coauthors, 2017: Lake volume and groundwater storage variations in Tibetan Plateau's endorheic basin. *Geophysical Research Letters*, **44**, 5550-5560.

Figures caption:

Figure 1. Analysis of near surface 2m air temperature (T2m) for boreal winter in 2008: (a) mean temperature, (b) location of gridboxes whose 6-hour T2m are less than a specific temperature threshold (TT) for the entire time during boreal winter.

Figure 2. Distribution of the northern hemisphere gauge undercatch correction factors, separately shown for rain and snowfall. The correction factors are based on Fuchs et al. (2001). The distributions are normalized so the total frequency for each phase is one.

Figure 3. Seasonal maps of two correction factors (CF): Legates climatology (CF-L) and Fuchs dynamic correction model (CF-F) shown in left and right columns, respectively. Maps of CF-F are calculated by averaging a decade (2000-2009) of monthly CFs for each season.

Figure 4. Seasonal maps of percent relative differences between CF-F and CF-L. The relative differences are calculated by subtracting CF-F from CF-L and dividing the result by the mean of the two. The calculation is performed using a decade (years 2000-2009) of monthly CF-Fs.

Figure 5. Comparison of summer- and winter-time mean precipitation of GPCP (5a and 5b) with different GPCC products (5c-5h) calculated from a decade of data (2003-2012). By multiplying CF-L and CF-F by GPCC Full data (Figs 5g and 5h), GPCC-L (Figs 5c and 5d) and GPCC-F (Figs 5e and 5f) are computed.

Figure 6. Maps of winter mean snowfall rate from GRACE, GPCP, GPCC, GPCC-F, GPCC-L, and CloudSat over frozen land in northcentral and northeast Asia. These maps are constructed at

2.5°x2.5° resolution using two years of data (2007-2008) to reduce potential noises. Corresponding maps of elevation (from sea surface) and station counts (from GPCC full data) are also provided.

Figure 7. Quantitative comparison of the products shown in Figure 6. The scatter plots are constructed from the maps displayed in Figure 8. Summary statistics are shown in each panel that include Correlation Coefficient (CC), Relative BIAS (relative to GRACE, so one is a perfect score), and Root Mean Square Error (RMSE) (mm day^{-1}).

Figure 8. Winter snowfall rates as a function of (a) T2m threshold and (b) surface elevation above sea level.

Figure 9. Time series of a decade (2004-2013) of winter mean snowfall rates from the studied products over frozen regions in north and northeast Asia. The mean snowfall rates are based on two-year running averages to enhance the stability of results. GRACE estimate is not shown for winter 2011, when more than a month of GRACE observations were unavailable due to battery management.

Figure 10. Quantitative statistics (CC, relative RMSE, and relative bias) calculated between GRACE winter snowfall rates and other products as a function of TT (top row) and elevation (bottom row) using a decade of corresponding datasets (2004-2013). Relative RMSE is calculated by dividing RMSE values by GRACE snowfall rate. Relative BIAS is defined as bias relative to GRACE, so one is a perfect score.

Figure 11. Similar to Figure 9, but is constructed over northwestern North America. The study area in 2008 is shown in Fig 1b.

Table 1. Quantitative comparison of GRACE with various products in estimating winter snowfall for 2003-2013 (except 2011). Comparison is made over northcentral and northeastern Asia (north of latitude 45°) for grid boxes that are frozen the entire time.

Products	CC	RMSE(mm d ⁻¹)	Relative bias
GPCP	0.63	0.51	1.09
GPCC	0.58	0.46	0.67
GPCC-L	0.62	0.55	1.10
GPCC-F	0.63	0.43	0.98

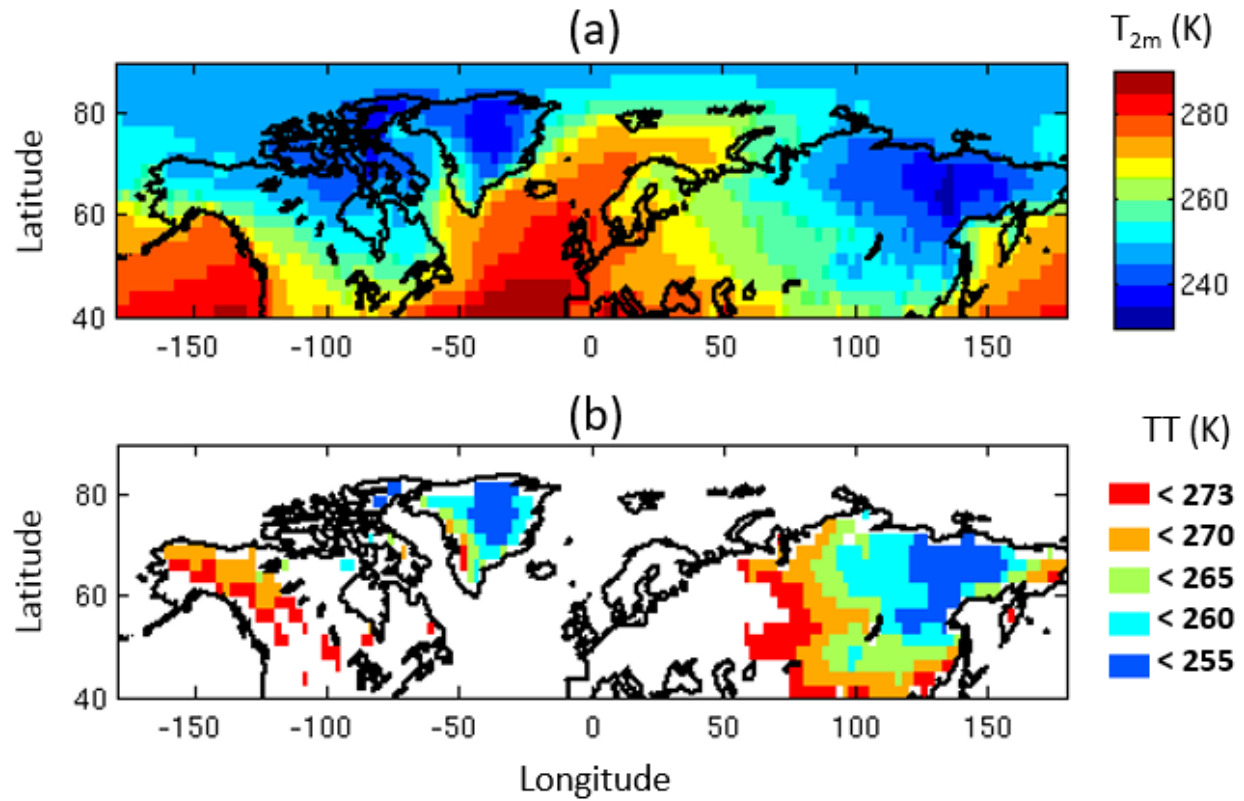


Figure 1. Analysis of near-surface 2m air temperature (T_{2m}) for boreal winter in 2008: (a) mean temperature, (b) location of gridboxes whose 6-hour T_{2m} are less than a specific temperature threshold (TT) for the entire time during boreal winter.

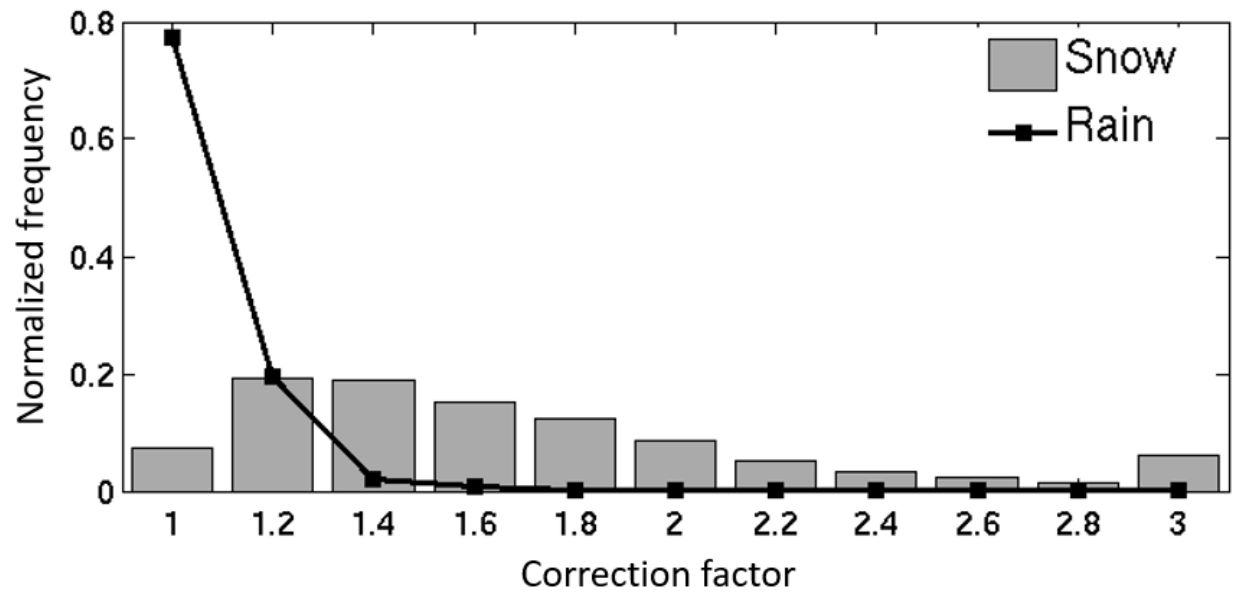


Figure 2. Distribution of the northern hemisphere gauge undercatch correction factors, separately shown for rain and snowfall. The correction factors are based on Fuchs et al. (2001). The distributions are normalized so the total frequency for each phase is one.

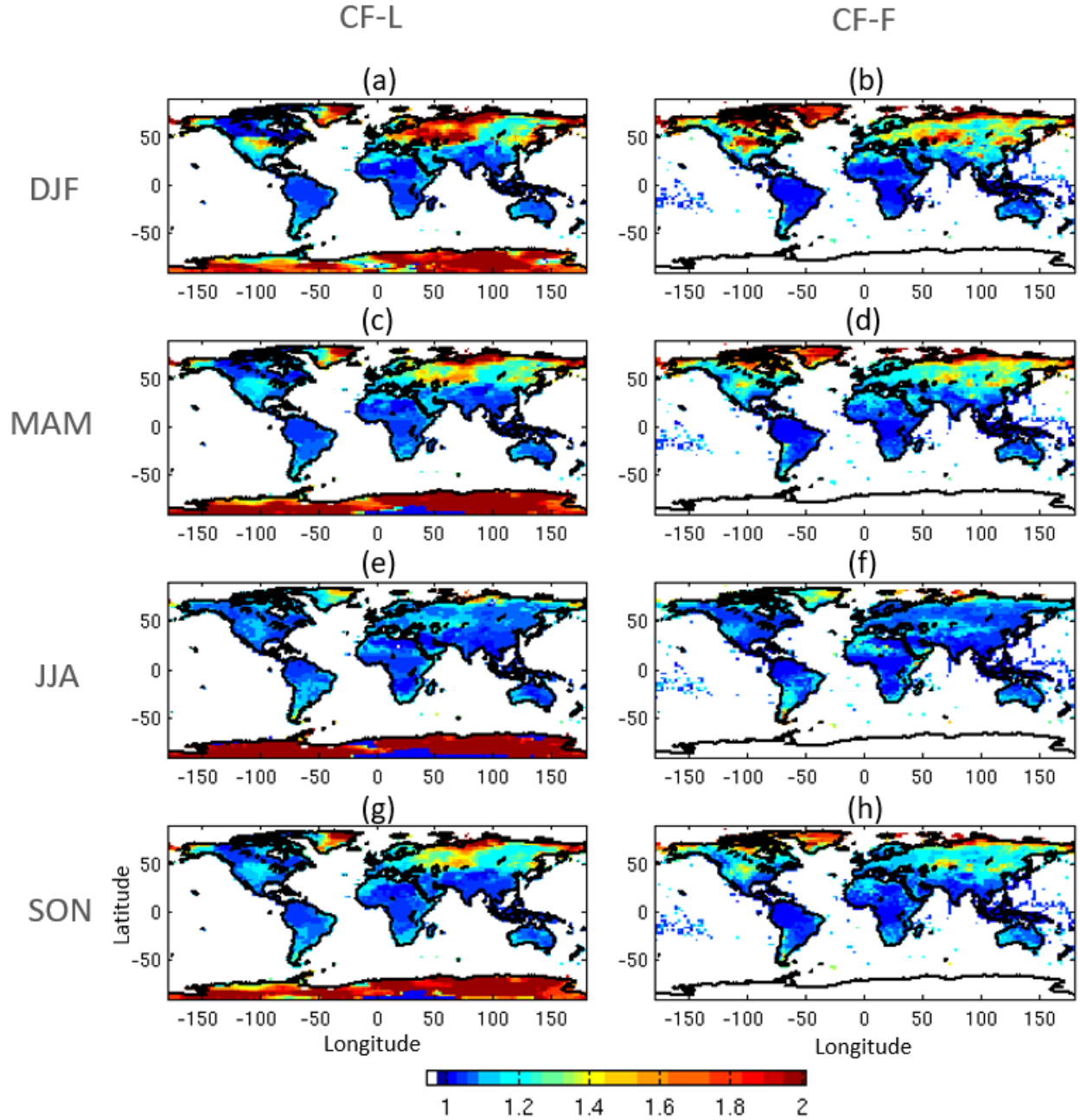


Figure 3. Seasonal maps of two correction factors (CF): Legates climatology (CF-L) and Fuchs dynamic correction model (CF-F) shown in left and right columns, respectively. Maps of CF-F are calculated by averaging a decade (2000-2009) of monthly CFs for each season.

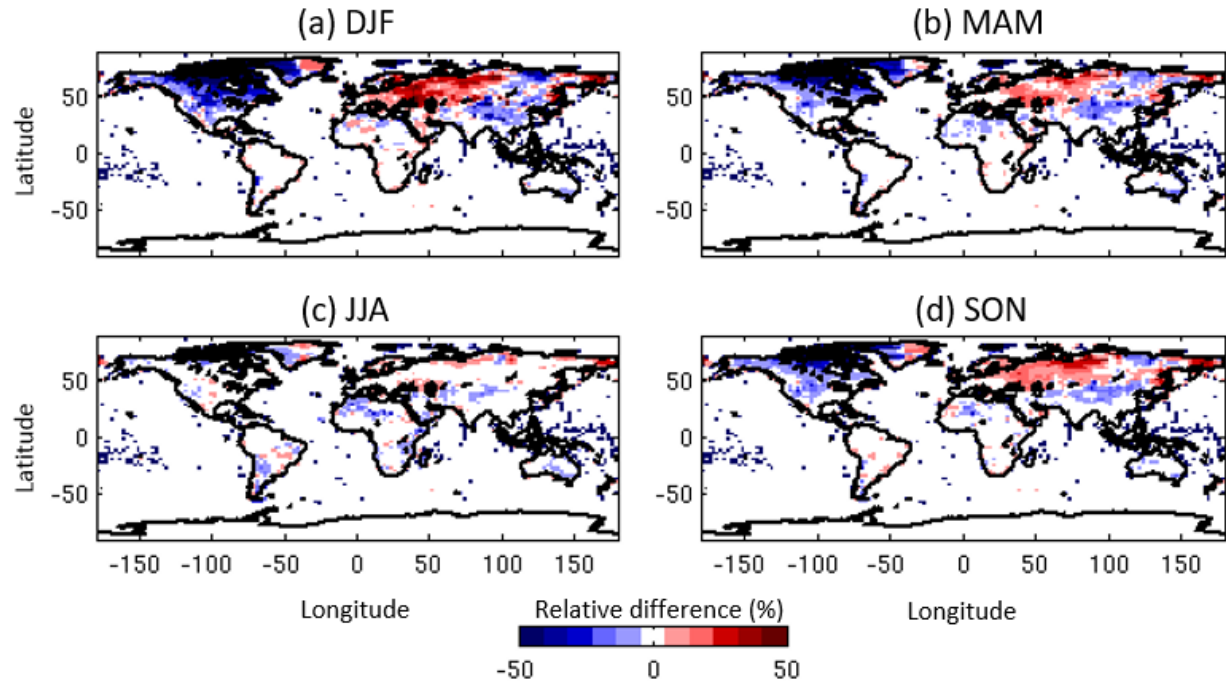


Figure 4. Seasonal maps of percent relative differences between CF-F and CF-L. The relative differences are calculated by subtracting CF-F from CF-L and dividing the result by the mean of the two. The calculation is performed using a decade (years 2000-2009) of monthly CF-Fs.

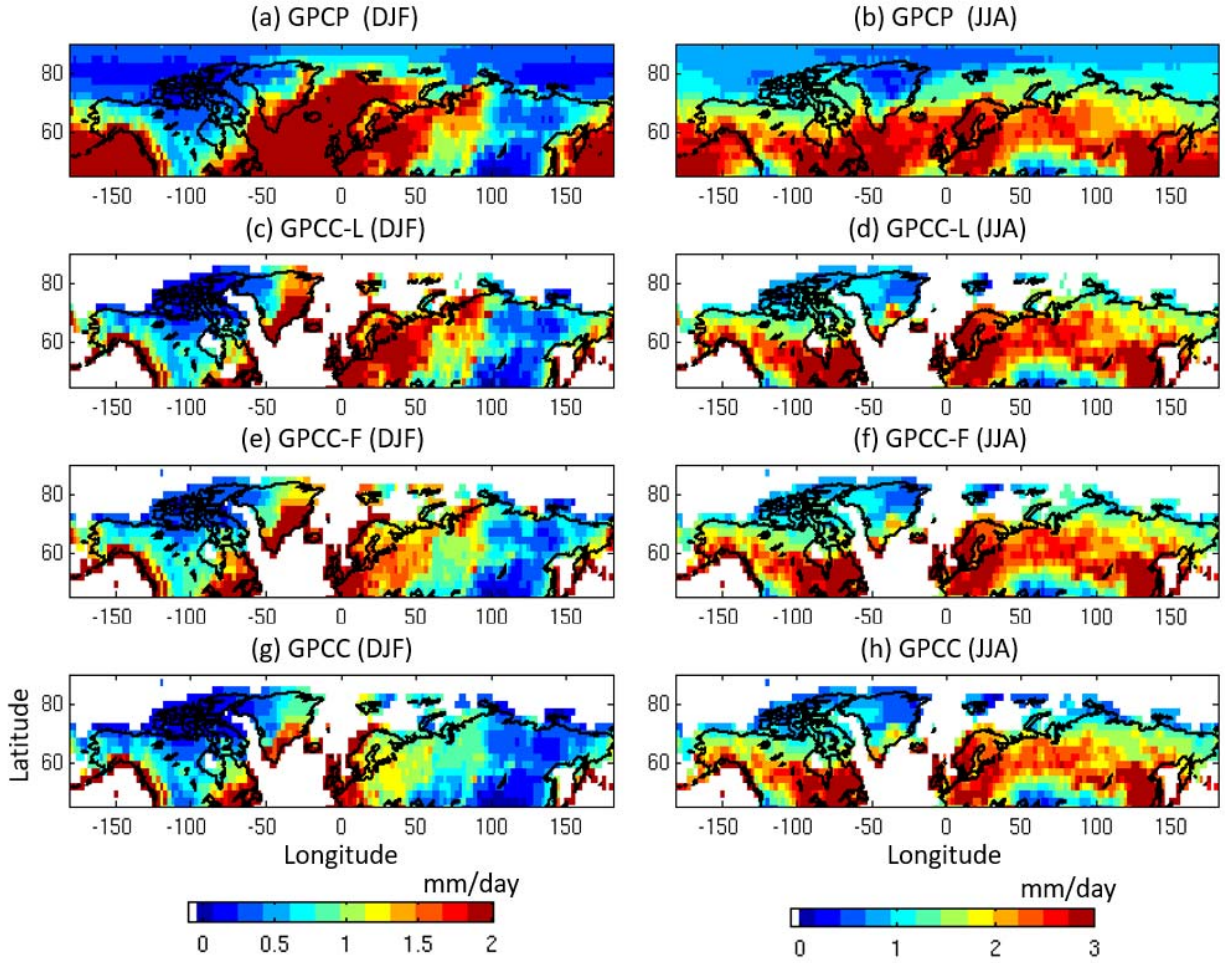


Figure 5. Comparison of summer- and winter-time mean precipitation of GPCP (5a and 5b) with different GPCC products (5c-5h) calculated from a decade of data (2003-2012). By multiplying CF-L and CF-F by GPCC Full data (Figs 5g and 5h), GPCC-L (Figs 5c and 5d) and GPCC-F (Figs 5e and 5f) are computed.

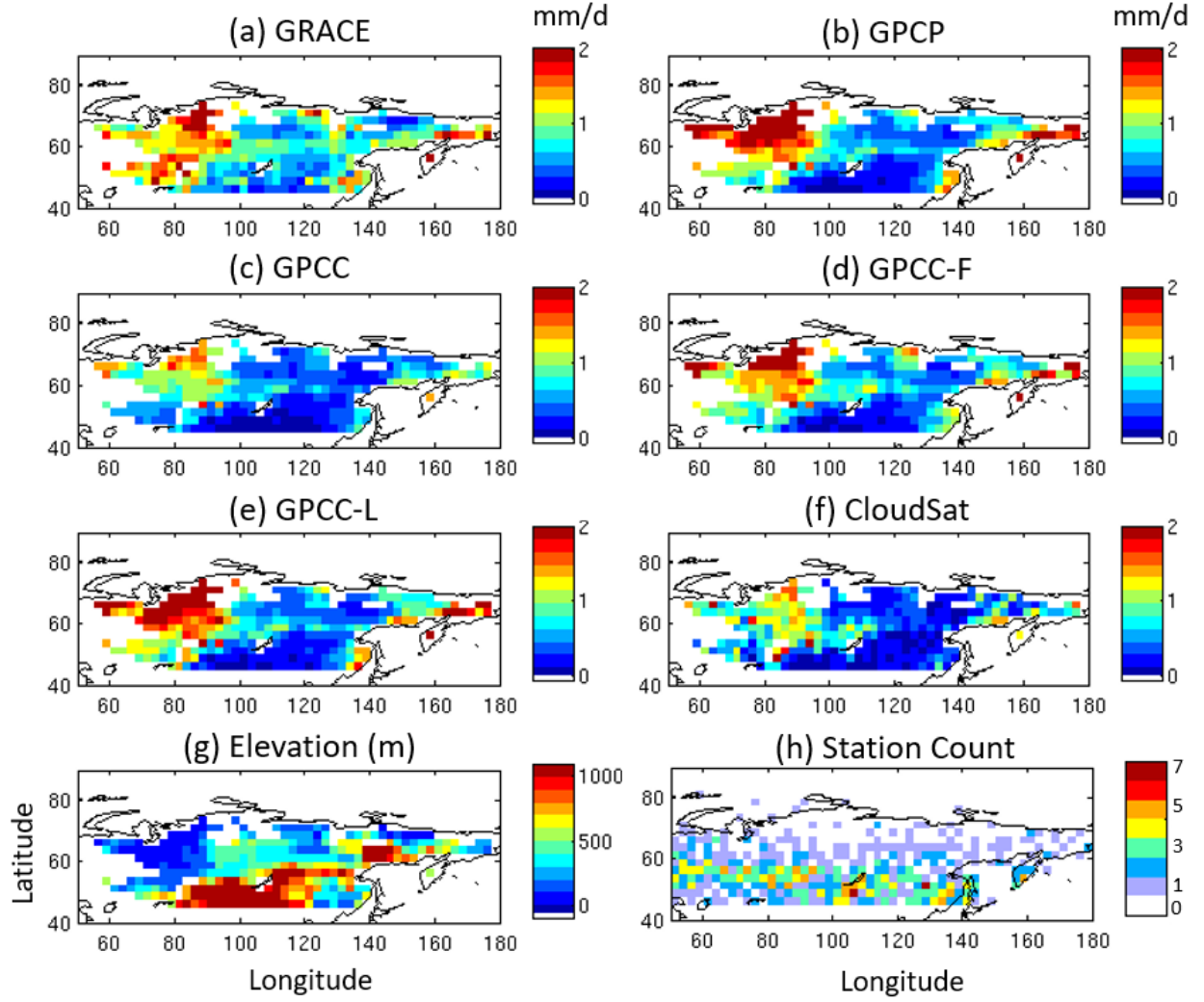


Figure 6. Maps of winter mean snowfall rate from GRACE, GPCP, GPCC, GPCC-F, GPCC-L, and CloudSat over frozen land in northcentral and northeast Asia. These maps are constructed at 2.5°x2.5° resolution using two years of data (2007-2008) to reduce potential noises. Corresponding maps of elevation (from sea surface) and station counts (from GPCC full data) are also provided.

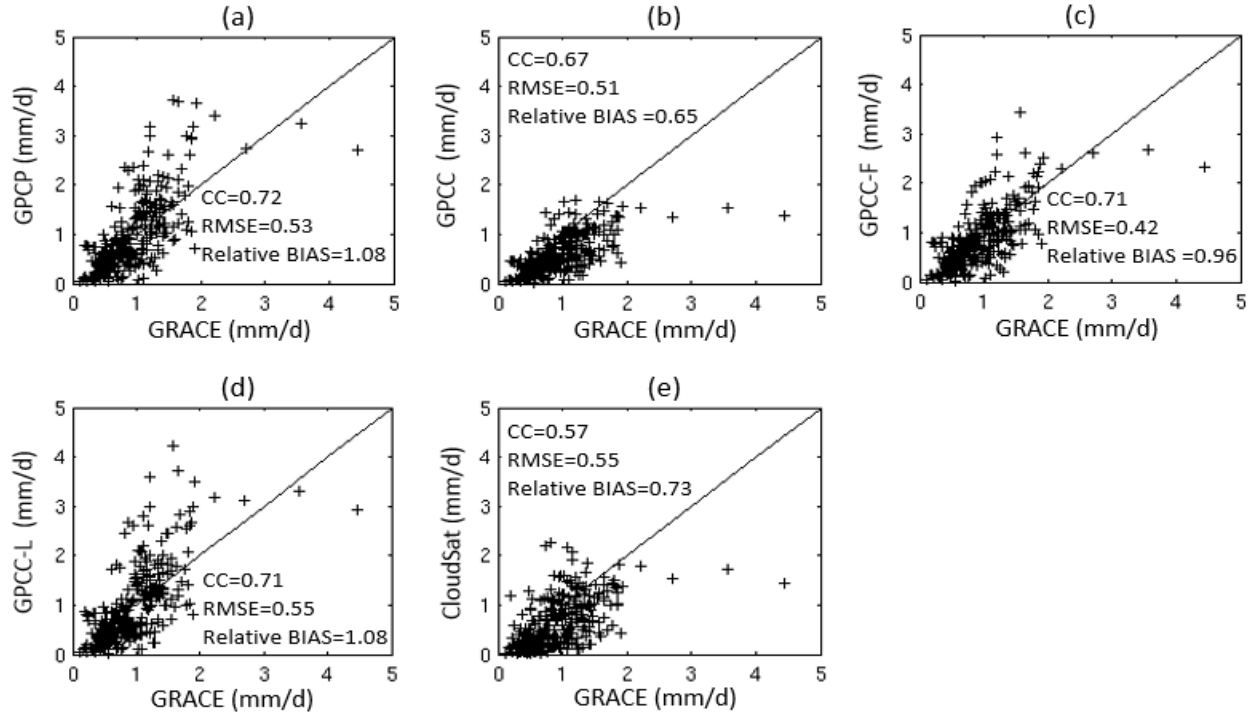


Figure 7. Quantitative comparison of the products shown in Figure 6. The scatter plots are constructed from the maps displayed in Figure 8. Summary statistics are shown in each panel that include Pearson Correlation Coefficient (CC), Relative BIAS (relative to GRACE, so one is a perfect score), and Root Mean Square Error (RMSE) (mm day^{-1}).

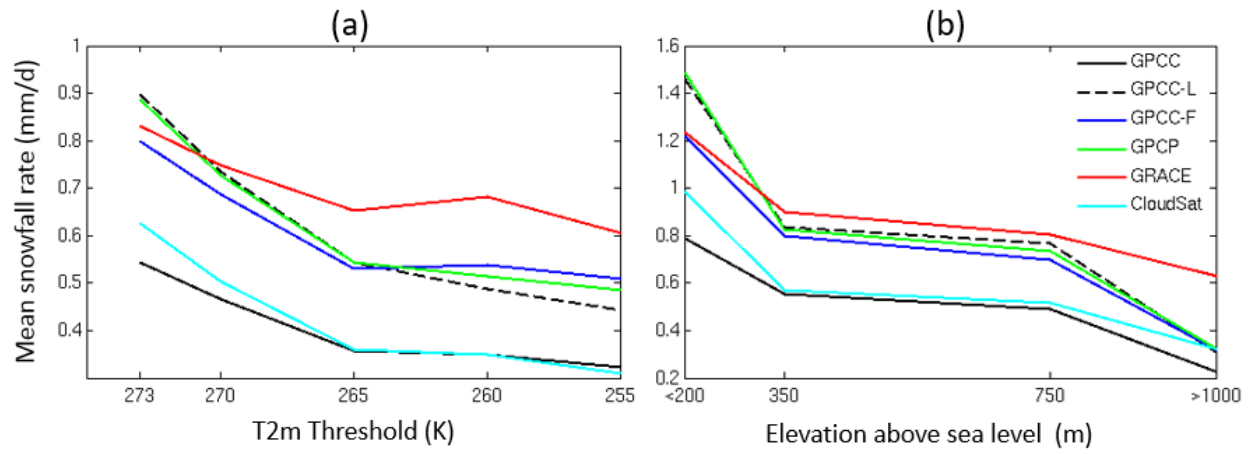


Figure 8. Winter snowfall rates as a function of (a) T2m threshold and (b) surface elevation above sea level.

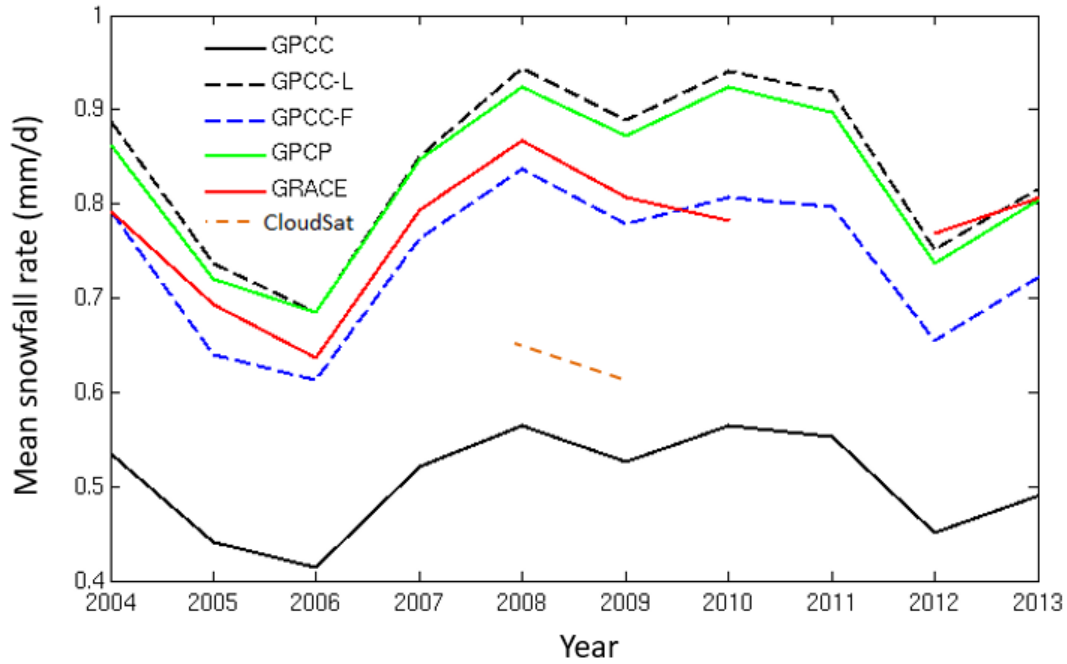


Figure 9. Time series of a decade (2004-2013) of winter mean snowfall rates from the studied products over frozen regions in north and northeast Asia. The mean snowfall rates are based on two-year running averages to enhance the stability of results. GRACE estimate is not shown for winter 2011, when more than a month of GRACE observations were unavailable due to battery management.

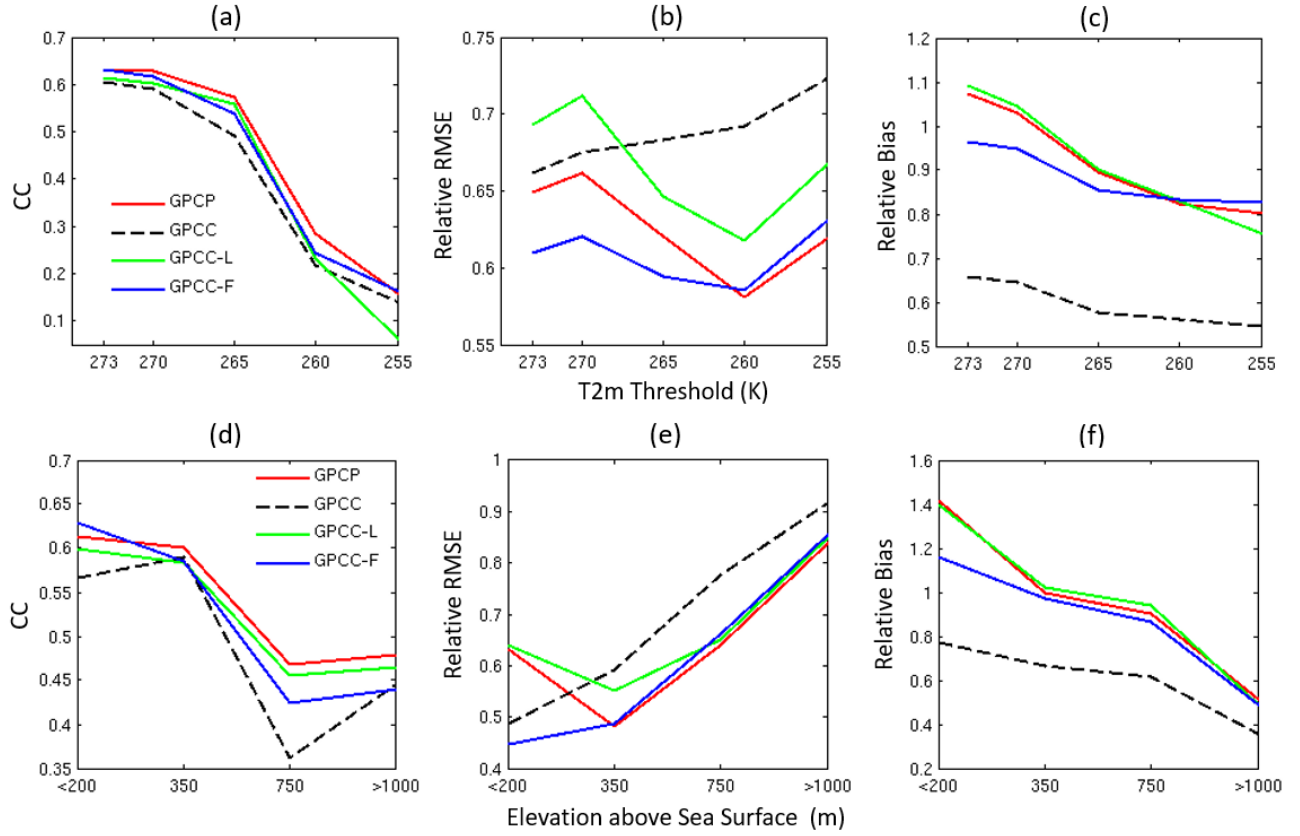


Figure 10. Quantitative statistics (CC, relative RMSE, and relative bias) calculated between GRACE winter snowfall rates and other products as a function of TT (top row) and elevation (bottom row) using a decade of corresponding datasets (2004-2013). Relative RMSE is calculated by dividing RMSE values by GRACE snowfall rate. Relative BIAS is defined as bias relative to GRACE, so one is a perfect score.

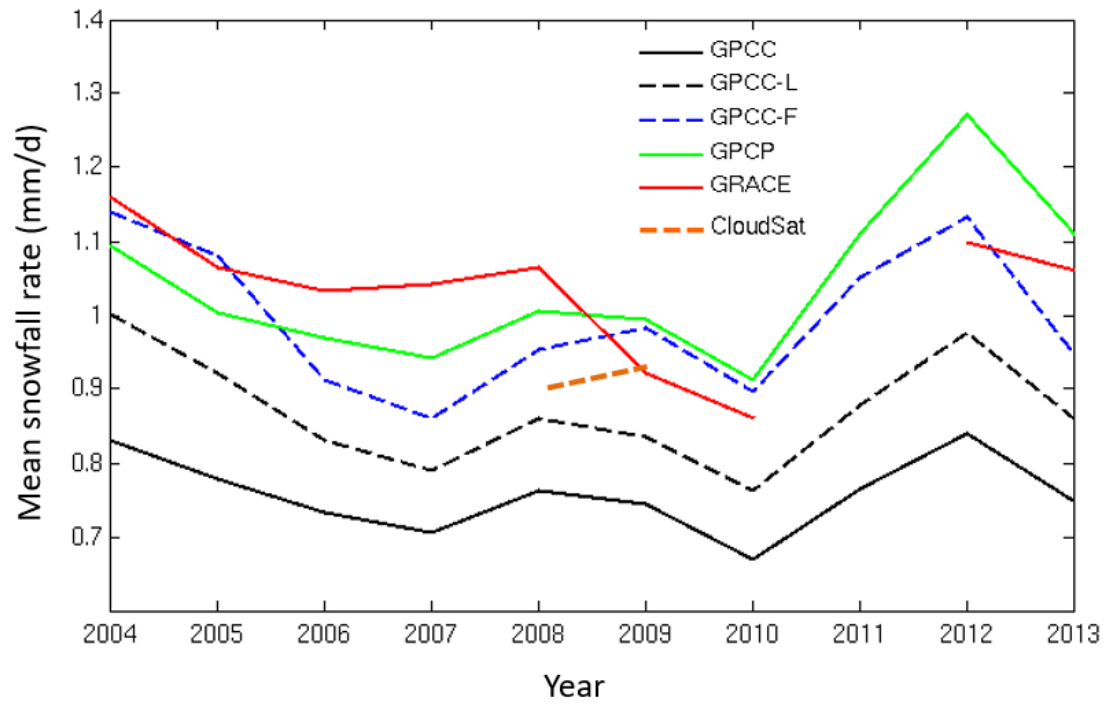


Figure 11. Similar to Figure 9, but is constructed over northwestern North America. The study area in 2008 is shown in Fig 1b.